CLAIMS: We claim:

1. A method for augmenting visual images of audio-visual entertainment systems, comprising the following steps of:

(a) enhancing one or a plurality of users' facial images in the video input, and

(b) simulating a virtual stage environment image

whereby the step for enhancing said user's facial image is processed in a facial image enhancement module automatically, dynamically and in real-time, and the step for simulating a virtual stage environment image is processed in a virtual stage simulation module.

2. The method according to claim 1, wherein the step (a) enhancing user's facial image in said video input further comprises a step for superimposing virtual object images to the facial image by a facial image enhancement process.

3. The method according to claim 1, wherein the step (a) enhancing user's facial image in said video input further comprises a step for superimposing virtual object images to the facial image by the embedded FET system in real-time.

4. The method according to claim 1, wherein the step (b) simulating a virtual stage environment image further comprises steps of:

(a) processing virtual object image selection,

(b) processing music selection, and

(c) composing virtual stage images.

5. The method according to claim 4, wherein the virtual stage simulation process further comprises a step for combining said user's enhanced face and body images with dynamically changing virtual background images,

whereby said virtual background images dynamically change according to said user's arbitrary movement.

6. The method according to claim 4, wherein the method further comprises a step for positioning a masked virtual stage image in front of said user's body image, whereby said user's body image is shown through the transparency channel region of said masked virtual stage image.

7. The method according to claim 4, wherein the method further comprises a step for using any arbitrary background rather than a controlled blue-screen style background for the actual background in the actual environment.

8. The method according to claim 4, wherein the method further comprises a step for using a standard background subtraction algorithm for said dynamic background construction.

9. The method according to claim 4, wherein the method further comprises a step for using said user's movement to trigger said dynamically changing virtual background images,

    whereby without said user's movement, said user's image could disappear behind said virtual background image,

    whereby this feature adds an interesting and amusing value to the system, in which said user has to dance as long as said user wants to see herself/himself on a means for displaying output, and

    whereby this feature can be utilized as a method for said user to participate in a dance in front of said audio-visual entertainment system.

10. The method according to claim 4, wherein the method further comprises a step for attaching musical instrument images, such as a guitar image or a violin image, to said user's body image,

    whereby said attached musical instrument images dynamically move along with said user's arbitrary motion in real-time, and

    whereby said user can also play the musical instrument by pretending as if he or she actually plays the musical instrument while looking at said musical instrument image on said means for displaying output.

11. An apparatus for augmenting visual images of an audio-visual entertainment system comprising:

(a) one or a plurality of means for capturing images,

(b) means for displaying output,

(c) means for processing and controlling,

(d) a sound system, and

(e) a microphone,

whereby the apparatus enhances one or a plurality of users' facial images in said captured images, automatically, dynamically and in real-time, and simulates a virtual stage environment image.

12. The apparatus according to claim 11, wherein the apparatus further comprises means for controlling lighting in a place where said system resides.

13. The apparatus according to claim 11, wherein said one or a plurality of means for capturing images further comprises one or a plurality of cameras and one or a plurality of frame grabbers.

14. The apparatus according to claim 11, wherein said one or a plurality of means for capturing images further comprises means for dynamically controlling fields of view.

15. The apparatus according to claim 11, wherein said one or a plurality of means for capturing images further comprises one or a plurality of Firewire or USB digital cameras.

16. The apparatus according to claim 11, wherein said means for processing and controlling further comprises means for performing the steps of:

    (a) enhancing facial images of said one or a plurality of users from said video input image sequences,

    (b) processing dynamically changing virtual background images according to said user's body movements,

    (c) simulating said virtual stage environment image by composing the enhanced facial and body image of said user, virtual stage images, and virtual objects images, and

    (d) handling interaction between said user and said system.

17. The apparatus according to claim 16, wherein the apparatus for step (a) enhancing facial images of said one or a plurality of users from said video input image sequences further comprises means for using a facial image enhancement process.

18. The apparatus according to claim 16, wherein the apparatus for step (a) enhancing facial images of said one or a plurality of users from said video input image sequences further comprises means for using the embedded FET system for said facial image enhancement process.

19. The apparatus according to claim 16, wherein the apparatus for step (c) simulating said virtual stage environment image further comprises means for preparing said virtual object images, such as musical instrument images and stage images, off-line.

20. The apparatus according to claim 11, wherein the apparatus further comprises means for using any arbitrary background as actual background for said user interaction.

21. A method for augmenting images on a means for displaying output of an audio-visual entertainment system, comprising the following steps of:

(a) capturing a plurality of images for an individual or a plurality of people with a single or a plurality of means for capturing images,

(b) processing a single image or a plurality of images from said captured plurality of images in order to obtain facial features and body movement information of said individual or said plurality of people,

(c) processing said user's selection for virtual object images on a means for displaying output,

(d) augmenting said user's facial image with said selected virtual object images,

(e) simulating a virtual stage environment image, and

(f) displaying said user's augmented facial image and said simulated virtual stage environment image on said means for displaying output,

whereby the step for augmenting said user's facial image with said selected virtual object images is processed automatically, dynamically and in real-time.

22. The method according to claim 21, wherein the method further comprises a step for processing touch-free interaction for said selection of said virtual object images.

23. The method according to claim 21, wherein the method further comprises a step for processing music selection.

24. The method according to claim 21, wherein the method further comprises a step for composing virtual stage images.

25. The method according to claim 21, wherein the method further comprises a step for combining said user's augmented facial and body image with dynamically changing virtual background images,
whereby said virtual background images dynamically change according to said user's arbitrary movement.

26. The method according to claim 25, wherein the method further comprises a step for positioning a masked virtual stage image in front of said user's body image,
whereby said user's body image is shown through the transparency channel region of said masked virtual stage image.

27. The method according to claim 25, wherein the method further comprises a step for using a standard background subtraction algorithm for said dynamic background construction.

28. The method according to claim 25, wherein the method further comprises a step for using said user's movement to trigger said dynamically changing background images,

    whereby without said user's movement, said user's image could disappear behind said background image,

    whereby this feature adds an interesting and amusing value to the system, in which said user has to dance as long as said user wants to see herself/himself on a means for displaying output, and

    whereby this feature can be utilized as a method for said user to participate in a dance in front of said audio-visual entertainment system.

29. The method according to claim 21, wherein the method further comprises a step for processing said captured plurality of images from any arbitrary background rather than a controlled background, such as a blue-screen,

    whereby said system can reside in any arbitrary environment, and

    whereby said processing of said single image or said plurality of images from said captured plurality of images in order to obtain facial features and body movement

information of said individual or said plurality of people, and said processing of said user's selection for virtual object images on a means for displaying output can be processed in said any arbitrary background.

30. The method according to claim 21, wherein the method further comprises a step for attaching musical instrument images, such as a guitar image or a violin image, to said user's body image,

whereby said attached musical instrument images dynamically move along with said user's arbitrary motion in real-time, and

whereby said user can also play the musical instrument by pretending as if he or she actually plays the musical instrument while looking at said musical instrument image on said means for displaying output.